

Secuenciación Genómica y Análisis del Genotipo Argentino de Soja Convencional “INTA-FICA 5C k/lx” [*Glycine max* (L.) Merr.]

Fumero M.V.^{1,2}, Bernardi C.¹, Garis S.B.^{1,2}, Lenzi L.¹, Amadio A.^{2,3}, Irazoqui M.^{2,3}, Cappuccio J.^{1,2}, Soldini D.¹, Vanzetti L.S.^{1,2,*}

¹Instituto Nacional de Tecnología Agropecuaria (INTA), EEA Marcos Juárez. Ruta 12 s/n, CP 2850 Marcos Juárez, Córdoba, Argentina.

²Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Godoy Cruz 2290, CP C1425FQB Buenos Aires, Argentina.

³Instituto Nacional de Tecnología Agropecuaria (INTA), EEA Rafaela. Ruta 34 Km 227, CP 2300 Rafaela, Santa Fe, Argentina.

*vanzetti.leonardo@inta.gob.ar

Introducción y Objetivos

La soja es un paleopoliploide y múltiples fuerzas evolutivas han dado forma a su genoma actual. En los últimos años, se han publicado múltiples ensamblados genómicos tanto para genotipos cultivados como silvestres, que han permitido comprender mejor la organización del genoma, detectar variabilidad y acelerar los procesos de mejoramiento. Pese a ser la principal oleaginosa producida y el tercer productor mundial de este cultivo, Argentina no posee antecedentes de secuenciación de variedades locales de sojas convencionales, no transgénicas, así como el desarrollo de herramientas genómicas aplicadas al mejoramiento genético en nuestro país es escaso. **El objetivo de este trabajo fue secuenciar y describir a nivel genómico a la variedad de soja convencional “INTA-FICA 5C k/lx”** (bajo inscripción en INASE), que incorpora características biológicas de calidad industrial diferencial, cuyo destino es la producción de harinas especiales y la alimentación de cerdos.

Resultados

- Luego de pasar el filtro de calidad mínima requerida ($Q>8$), se obtuvieron 4,886,929 lecturas de las cuales el 78.7% mapearon de manera homogénea a lo largo de los 20 cromosomas de *G. max* William 82 (Figura 1)
- La longitud promedio de las lecturas fue de 1,335pb, abarcado un rango entre 500pb y 47,000pb (Figura 1)
- Se detectaron 1,154,205 variantes homogéneamente distribuidas a lo largo de los cromosomas, en una proporción de 1/1000bp, excepto en los cromosomas 3 y 16 en los cuales fue de 2 y 3/1000pb (Figura 2)
- El 87.7% de las variantes se ubicaron en regiones intergénicas, mientras que el restante 12.3% (142,362) fueron detectadas sobre genes (Figura 2)
- Los cromosomas 3 y 16 exhibieron la mayor cantidad de variantes (Figura 3)

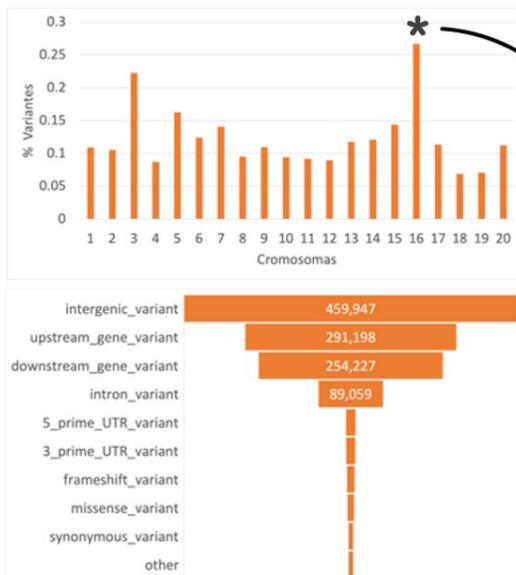


Figura 2. Proporción de variantes por cromosoma (arriba) y tipos de variantes según posición genómica (abajo)

Metodología y Análisis Bioinformáticos

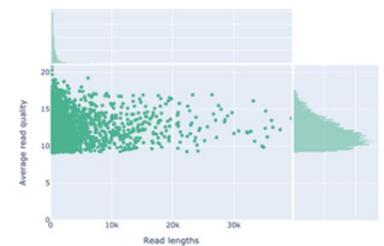
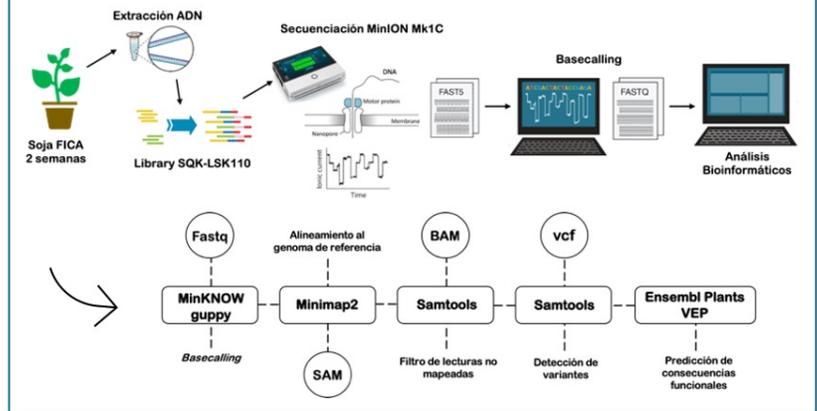


Figura 1. Longitud de las lecturas versus calidad promedio

FICA - Cromosoma 16

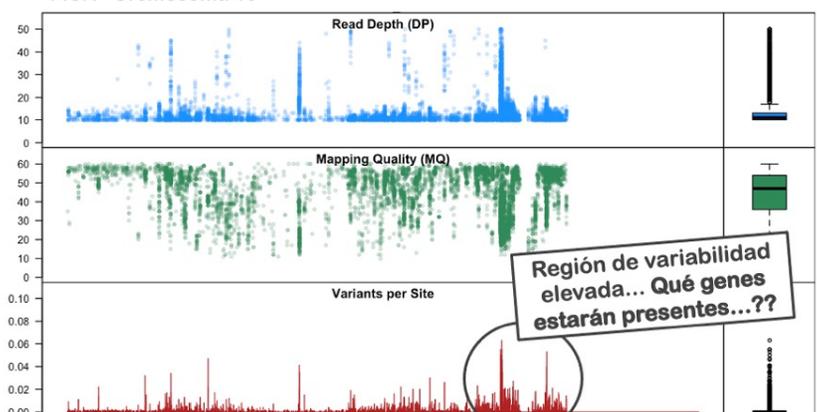


Figura 3. Sliding-window mostrando la distribución de profundidad y calidad del mapeo y la cantidad de variantes por sitio sobre el crom.16, en ventanas de 1000pb

Conclusión y Perspectivas. Poseer las secuencias genómicas de variedades o genotipos de élite adaptados abrirán posibilidades a la implementación de nuevas tecnologías como la edición sobre genes de interés de una manera eficiente y podrán contribuir al desarrollo y aplicación de *new breeding technologies* (NBTs) en los programas de mejoramiento clásicos